NASA Technical Memorandum 100649

TWO BIASED ESTIMATION TECHNIQUES IN LINEAR REGRESSION--APPLICATION TO AIRCRAFT

(NASA-TM-100649) TWO BIASED ESTIMATION TECHNIQUES IN LINEAR REGRESSION: APPLICATION TO AIRCRAFT (George Washington Univ.) 41 p
CSCL 12A
N88-29489
Unclas
G3/65 0158987

VLADISLAV KLEIN

JULY 1988

National Aeronautics and Space Administration

Langley Research Center
Hampton, Virginia 23665

SUMMARY

The report is an introduction to an important but relatively neglected aspect of regression theory which deals with near linear dependency in measured data, called collinearity. Several ways for detection and assessment of collinearity are discussed. Because data collinearity usually results in poor least squares estimates, two estimation techniques which can limit a damaging effect of collinearity are presented. These two techniques, the principal components regression and mixed estimation, belong to a class of biased estimation techniques.

Data collinearity detection and assessment, and the two biased estimation techniques are demonstrated in two examples using flight test data from longitudinal maneuvers of an experimental aircraft. The eigensystem analysis and parameter variance decomposition appeared to be a promising tool for collinearity evaluation. The biased estimators had far better accuracy than the results from the ordinary least squares technique.

SYMBOLS AND ABBREVIATIONS

A	matrix of known constants
a	vector of specified values
a_z	vertical acceleration, g units
b	unknown vector
C_Z	vertical-force coefficient
C_m	pitching-moment coefficient
Cov(•)	covariance matrix
c	constant
$\bar{c}$	mean aerodynamic chord, m
D	diagonal matrix
d	constant indicating a region for a priori value
E (•)	expected value
g	acceleration due to gravity, m/sec^2
I	identity matrix
I_Y	moment of inertia about lateral body axes, $kg-m^2$
k	constant
M	matrix defined by eq. (37)
MSE (•)	mean square error
m	mass, kg
N	number of data points
n	number of regressors
p	number of prior restrictions on elements of θ
q	pitch rate, rad/sec or deg/sec
$\bar{q}$	dynamic pressure, $\rho V^2/2$, Pa
R^2	squared multiple correrlation coefficient

R_j^2	squared multiple correlation coefficient of x_j regressed on the remaining regressors
r	rank of X matrix
S	wing area
S_j	square root of sum of squared differences $x_{ji} - \bar{x}_j$
s	n - r
s (•)	standard error
T	matrix of eigenvectors
t	time, sec
t_j	jth column of T matrix
U	orthogonal matrix
V	airspeed, m/sec
VIF	variance inflation factor
W	weighting matrix
X	matrix of regressors and ones
x	regressor
Y	vector of dependent variables
y	dependent variable
Z	matrix of orthogonal regressors
α	angle of attack, rad or deg
γ	parameters in model with orthogonal regressors
$\delta_c, \delta_f, \delta_s$	canard, flaperons, and strake deflection respectively, rad or deg
ε	vector of measurement noise
ζ	random vector
η	condition index
Θ	vector of unknown parameters (regression coefficients)

κ	condition number
Λ	matrix of eigenvalues
λ	eigenvalue of X^TX
μ	singular value of X
ξ	sensitivity
π_{kj}	variance proportion of jth regression coefficient associated with kth component of its decomposition
ρ	air density, kg/m^3
σ^2	variance

Superscripts:

$\hat{}$	least squared estimate
$\sim$	biased estimate
*	standardized regressors
,	scaled regressors
•	derivative with respect to time

Matrix Exponents:

T	transpose matrix
-1	inverse matrix

Abbreviations:

LS	least squares
ME	mixed estimation
PC	principal components
SVD	singular value decomposition

INTRODUCTION

Recently, the introduction of highly maneuverable and often inherently unstable aircraft has been presenting new challenges to aircraft identification and parameter estimation. These new aircraft may have more control surfaces than conventional aircraft which are moved through a flight control system. Such a system can introduce a close relationship between the deflections of various surfaces and at the same time can preclude maneuvers suitable for system identification. These characteristics can be reflected in an inability to estimate the effectiveness of individual control surfaces and to obtain accurate estimates of the remaining parameters. One of the reasons for these problems is related to the near linear relationship among several variables entering model for various parameter estimation techniques.

Near linear dependency among variables in linear regression, often called collinearity, has been studied by many statisticians. An introduction to the problem of collinearity is presented in ref. 1. The purpose of this report is to briefly discuss the collinearity in a general model for linear regression, detection of collinearity and its remedy. Two methods of dealing with collinear data, the principal components regression and mixed estimation, are presented. They are based on an extension of the ordinary least squares technique. The report is concluded by two examples with real flight data. In these examples the detection of collinearity and application of estimation techniques described is demonstrated.

COLLINEARITY

The linear regression model can be formulated as

$$y = \theta_0 + \theta_1 x_1 + \ldots + \theta_n x_n \tag{1}$$

where x_j, $j = 1, 2, \ldots, n$, are the regressors, y is a dependent variable and $\theta_0, \theta_1, \ldots, \theta_n$ are the unknown parameters. After substituting measured values into (1) the regression equation has the form

$$Y = X\theta + \varepsilon \tag{2}$$

where Y is an (N x 1) and θ is (n + 1 x 1) vector, ε is an (N x 1) vector of measurement noise and X is the (N x n + 1) matrix of regressors and ones

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} \cdots x_{n1} \\ 1 & x_{12} & x_{22} \cdots x_{n2} \\ \vdots & \vdots & \vdots \\ 1 & x_{1N} & x_{2N} \cdots x_{nN} \end{bmatrix}$$

with N indicating the number of data points. The least squares estimates of unknown parameters are obtained from

$$\hat{\theta}_{LS} = (X^T X)^{-1} X^T Y \tag{3}$$

For further discussion and analysis it will be more convenient to deal with regressor variables which have been standardized (centered and scaled to unit length), see Appendix A. There, the matrix $X^{*T}X^*$ is the (n x n) matrix matrix of correlations because the off-diagonal elements of this matrix are quite often referred to as correlation coefficients, although the regressors are not necessarily random variables. Denoting X^*_j, $j = 1, 2, \ldots, n$, as the columns of the X* matrix with centered and scaled regressors, the matrix X* can be expressed as

$$X^* = [X^*_1, X^*_2, \ldots, X^*_n] \tag{4}$$

If $X^{*T}_j X^*_k = 0$, $j \neq k$, the regressors are orthogonal and the $X^{*T}X^*$ matrix is a diagonal matrix. The vectors $X^*_1, X^*_2, \ldots, X^*_n$ are called linearly dependent if there is a set of constants, k_j, not all zero, such that

$$\sum_{j=1}^{n} k_j X^*_j = 0 \tag{5}$$

Then, the rank of $X^{*T}X^*$ is less than n and $(X^{*T}X^*)^{-1}$ does not exist.

In many practical applications of linear regression eq. (5) is only approximately true. This indicates near linear dependency in X^* and the problem of collinearity exists. In such case $X^{*T}X^*$ is called ill conditioned. Because of that collinearity can cause computational problems and reduce the accuracy of estimates. Thus in the context of linear regression, collinearity is a data problem, not a statistical phenomenon.

There are at least three different sources of collinearity, namely,

a) design of an experiment,
b) constraints in the data,
c) model specification.

If the model is designed in such a way that the resulting data is specified mostly on a subspace of the region defined approximately by (5), then collinearity might occur. This type of problem can arise during the test of a dynamical system where one or more variables representing regressors in (1) were not sufficiently excited. The constraints in the data could be caused by an inherent property of the system under test. For example, an aircraft stability augmentation system can deflect various control surfaces in concert thus causing near linear dependence among their deflections. Finally, to avoid collinearity, a specified model should not be over parameterized. For example, it should not include nonlinear terms, such as x_1^2, or $x_1 x_2$, if x_1 is small.

The presence of collinearity usually results in various unwanted properties of the least squares estimates of unknown parameters. Two of them, illustrated in ref. 1, include too large absolute values for parameter estimates and their large variances and covariances.

DETECTION AND ASSESSMENT OF COLLINEARITY

Many procedures have been employed to detect collinearity. They are discussed in ref. 1 and ref. 2. In this report only three of them will be considered and their use later demonstrated in examples. These procedures are:

1. Examination of the correlation matrix and its inverse.

This is the simplest and more straight forward procedure. High correlation coefficient between two regressors can point to a possible collinearity problem. The absence of high correlation, however, cannot be viewed as evidence of no problem. The correlation matrix is unable to reveal the presence of several coexisting near dependencies among the regressors, as demonstrated in ref. 1. Because of the shortcoming mentioned in regard to the use of $X^{*T}X^{*}$ as a diagnostic measure of collinearity, the usefulness of its inverse is also limited. The diagonal elements $(X^{*T}X^{*})^{-1}$ are often called the variance inflation factors, VIF_j, and they can be expressed as

$$VIF_j = \frac{1}{1 - R_j^2} \tag{6}$$

where R_j^2 is the squared multiple correlation coefficient of x_j regressed on the remaining regressors (see ref. 1 and Appendix B for the development of R^2). The term "variance inflation factor" reflects its relationship with the jth parameter variance $\sigma^2(0_j)$. As shown in ref. 3

$$\sigma^2(0_j) = \frac{\sigma^2}{X_j^{*T}X_j^{*}} VIF_j \tag{7}$$

The diagnostic value of VIF follows from expression (6). Large value of VIF indicates an R_j^2 near unity and hence points to collinearity. The weakness of this diagnostic measure is in its inability to distinguish among several coexisting near dependencies and in its lack of meaningfull boundaries for values of VIF.

2. Eigensystem Analysis and Singular Value Decomposition

The matrix X^TX can be decomposed as

$$X^TX = T\Lambda T^T \tag{8}$$

where Λ is an (n x n) diagonal matrix whose diagonal elements are the eigenvalues λ_j, j = 1, 2, . . ., n, of X^TX, and T is an (n x n) orthogonal matrix

whose columns are the eigenvectors of X^TX. The eigenvalues close to zero indicate near linear dependency in the data. The elements of the corresponding eigenvectors could reveal the nature of this dependency. Collinearity is, therefore, indicated by the presence of a "small" eigenvalue. Unfortunately there is no specification what "small" is. In order to avoid this problem some authors are using the condition number of X^TX defined as

$$\kappa_j = \frac{\lambda_{max}}{\lambda_j}, \quad j = 1, 2, \ldots, n \tag{9}$$

Then, they consider the condition number exceeding 1000 as an indication of severe collinearity (see ref. 1).

In ref. 2 an approach using singular-value decomposition for diagnosing collinearity is recommended. It is based on the decomposition of matrix X as

$$X = UDT^T \tag{10}$$

where U is a (N x n) matrix and $U^TU = T^TT = I$. The matrix D is an (n x n) diagonal matrix with nonnegative diagonal elements μ_j, $j = 1, 2, \ldots, n$, which are called the singular values of X. The singular-value decomposition (SVD) is closely related to the concept of enigenvalues and eigenvectors, since from (8) and (10)

$$X^TX = TD^2T^T = T\Lambda T^T \tag{11}$$

The diagonal elements of D^2 are therefore the eigenvalues of X^TX and the columns of U are the eigenvectors of X^TX associated with its n nonzero eigenvalues. The degree of ill conditioning depends on how small the singular value is relative to the maximum singular value. In this connection a condition index of the matrix X is proposed as

$$\eta_j = \frac{\mu_{max}}{\mu_j}, \quad j = 1, 2, \ldots, n \tag{12}$$

It is further suggested to consider η_j from 30 to 100 as an evidence of moderately to strongly collinear data.

The SVD of the matrix X provides similar information to that given by the eigensystem of X^TX. The use of SVD is, however, preferred by many authors namely because of greater numerical stability of its computing in comparison to that of the eigensystem of X^TX. This may be especially true when X^TX is ill conditioned.

3. Parameter Variance Decomposition

The parameter variance decomposition approach for detecting collinearity was proposed in ref. 2. It follows from the covariance matrix of parameter estimates $\hat{\theta}$ which is obtained as

$$\text{Cov}\ (\hat{\theta}) = \sigma^2(X^TX)^{-1} = \sigma^2\ T\Lambda^{-1}T^T \tag{13}$$

The variance of each parameter is equal to

$$\sigma^2(\hat{\theta}_j) = \sigma^2 \sum_{k=1}^{n} \frac{t_{jk}^2}{\lambda_j} = \sigma^2 \sum_{k=1}^{n} \frac{t_{jk}^2}{\mu_j^2} \tag{14}$$

where t_{jk} are the elements of eigenvector t_j associated with λ_j. Eq. (14) decomposes the variance of each parameter into a sum of components, each corresponding to one and only one of the n singular values μ_j. In (14) the singular values appear in denominator, so one or more small singular values can substantially increase the variance of θ_j. This means that an unusually high proportion of the variance of two or more coefficients for the same small singular value can provide evidence that the corresponding near dependency is causing problems. Introducing

$$\phi_{jk} = \frac{t_{jk}^2}{\mu_j^2} \qquad \text{and} \qquad \phi_j = \sum_{k=1}^{n} \phi_{jk}\ ,$$

the j,k variance-decomposition proportion as the proportion of the variance of the jth regression coefficient associated with the kth components of its decomposition in (14) is given as

$$\pi_{kj} = \frac{\phi_{jk}}{\phi_j},\ j,\ k = 1,\ 2,\ \ldots,\ n \tag{15}$$

Since two or more regressors are required to create near dependency, then two or more variances will be adversely affected by high variance-decomposition proportions associated with a single singular value. Variance-decomposition proportions greater than 0.5 are recommended in ref. 2 as a guidance for possible collinearity problems. It is also suggested that the columns of X should be scaled to unit length but not centered. Thus the role of the bias term in near-linear dependencies can be diagnosed.

SENSITIVITY ANALYSIS

As was mentioned earlier, the design of an experiment and constraint in the data can contribute to data collinearity. Both of these phenomena may also influence parameter identifiability resulting in limited accuracy of their estimates. One of the possible ways to assess parameter identifiability is based on the sensitivity analysis. For the regression model

$$Y = X\Theta + \varepsilon \tag{2}$$

the sensitivity of the dependent variables to the changes in parameter Θ_j, keeping the remaining parameters fixed, is given as

$$\frac{\partial Y}{\partial \Theta_j} = X \frac{\partial \Theta}{\partial \Theta_j} \tag{16}$$

Then, the measure of sensitivity for the parameter Θ_j can be defined as

$$\xi_j = \Theta_j^2 \left(\frac{\partial Y}{\partial \Theta_j}\right)^T \left(\frac{\partial Y}{\partial \Theta_j}\right) = \Theta_j^2 \left(\frac{\partial \Theta}{\partial \Theta_j}\right)^T X^T X \left(\frac{\partial \Theta}{\partial \Theta_j}\right) = \Theta_j^2 X_j^T X_j \tag{17}$$

For practical computing of the sensitivities the values of parameters in the regression model must be known. Because the parameter values are the subject of estimation, the question can arise what values for Θ_j should be used in computing ξ_j. The least squares estimates using data with strong collinearity and/or low parameter sensitivity could be highly unstable thus causing distortions in the computed values of ξ_j. In these cases more stable estimates or priori values for parameters should be used.

BIASED ESTIMATION TECHNIQUES

There are several ways on how to deal with the problem of collinearity. They include a collection of additional data, redesign of an experiment, model respecification, and use of different estimation techniques from the ordinary least squares procedure. This report will address only the last possibility mentioned.

As discussed previously, the application of the ordinary least squares technique to the set of data with collinearity problems can result in large estimated values for parameters and large values for their covariances. The

least squares technique provides an unbiased linear estimator which, according to Gauss-Markoff theorem (see for example ref. 4), has minimum variance in the class of unbiased linear estimators. There is no guarantee, however, that this variance will be small. Figure 1 illustrates a situation of two distributions of a parameter estimate. One estimate, $\hat{\theta}$, is unbiased (a possible result of least squares technique), the other, $\tilde{\theta}$, is biased (obtained by a biased estimation technique). In the first case the variance of $\hat{\theta}$ is large, indicating a large confidence interval on θ and unstable point estimate $\hat{\theta}$. In the second case the estimate $\tilde{\theta}$ is subjected to bias error, $E(\tilde{\theta}) - \theta$, but much smaller variance. The resulting mean square error of the estimator $\tilde{\theta}$ is

$$MSE(\tilde{\theta}) = E(\tilde{\theta} - \theta)^2 = \sigma^2(\tilde{\theta}) + [E(\tilde{\theta}) - \theta]^2 \tag{18}$$

It is possible that for small bias error the MSE $(\tilde{\theta})$ could be smaller than the variance of the least squares estimator $\sigma^2(\hat{\theta})$. This possibility has inspired a development of various biased estimation techniques. Two of them, the principal components regression and mixed estimation, will be described and applied to experimental data.

PRINCIPAL COMPONENTS REGRESSION

The development of principal components estimator starts by transforming the original regressors x_j, $j = 1, 2, \ldots, n$ to the space of orthogonal regressor z_j. This transformation is accomplished by introducing

$$Z = XT \tag{19}$$

and

$$\theta = T\gamma \tag{20}$$

where

$$T^T X^T X T = \Lambda \text{ and } T^T T = TT^T = I$$

Using (19) and (20) the regression model (2) becomes

$$Y = Z\,\gamma + \varepsilon \tag{21}$$

with the LS estimator of γ as

$$\hat{\gamma} = \Lambda^{-1} Z^T Y \tag{22}$$

The columns of Z which define a new set of orthogonal regressors are referred to as principal components.

To obtain principal components estimator the regressors in (21) are arranged in order of decreasing eigenvalues

$$\lambda_1 > \lambda_2 > \cdots > \lambda_n$$

Then, the principal component estimator is given by a vector where the first r components agree with $\hat{\gamma}$ and remaining s = n - r components are zero

$$\tilde{\gamma}_{PC} = \left[\hat{\gamma}_1, \hat{\gamma}_2, \ldots, \hat{\gamma}_r, 0 \ldots 0\right]^T \tag{23}$$

The LS estimates in (22) can also be obtained as

$$\hat{\gamma}_j = \lambda_j^{-1}\, t_j^T X^T Y, \quad j = 1, 2, \ldots, n \tag{24}$$

where t_j is the jth column of the eigenvector matrix T. By comparing (23) and (24) it follows that the estimates $\tilde{\gamma}_{PC}$ are obtained by setting s = n - r small eigenvalues to zero which is equivalent to assuming that the matrix X has rank r < n.

The principal components estimates of parameters associated with the original regressors x_j are obtained from (20) and (23) as

$$\tilde{\theta}_{PC} = \left[t_1, t_2, \ldots, t_n\right] \tilde{\gamma}_{PC} \tag{25}$$

In order to find the expression for the bias in principal components estimates and variance of these estimates the eigenvlaue and eigenvector matrices are partitioned as

$$\Lambda = \begin{bmatrix} \Lambda_r & 0 \\ 0 & \Lambda_s \end{bmatrix}$$

$$T = [T_r \quad T_s]$$

where Λ_r and Λ_s are diagonal matrices containing the eigenvalues associated with the retained and eliminated principal components respectively. For the eigenvectors, T_r and T_s are similarly partitioned. The LS estimator of the parameters γ that are retained is

$$\hat{\gamma}_r = \Lambda_r T_r^T X^T Y \tag{26}$$

From (24), (25) and (26) it follows that

$$\tilde{\theta}_{PC} = T_r \hat{\gamma}_r = \sum_{j=1}^{r} \lambda_j^{-1} t_j^T X^T Y t_j \tag{27}$$

The expected value of the PC estimator is

$$E(\tilde{\theta}_{PC}) = T_r \gamma_r = T_r T_r^T \theta$$

Since $$TT^T = I = T_r T_r^T + T_s T_s^T$$

$$E(\tilde{\theta}_{PC}) = [I - T_s T_s^T]\,\theta$$

$$= \theta - T_s \gamma_s$$

Thus the PC estimates of the n parameters θ are biased by the quantity $T_s \gamma_s$.

Assuming that ε has zero mean and variance $\sigma^2 I$, the covariance matrix of the LS estimates of θ is given as

$$\text{Cov}(\hat{\theta}) = \sigma^2 (X^T X)^{-1} = \sigma^2 T \Lambda^{-1} T^T \tag{28}$$

$$= \sigma^2 (T_r \Lambda_r^{-1} T_r^T + T_s \Lambda_s^{-1} T_s^T)$$

The form of the covariance matrix for the PC estimates of 0 follows from (13), (14) and (27) as

$$\text{Cov}\,(\tilde{\theta}_{pc}) = \sigma^2 T_r \Lambda_r^{-1} T_r^T + \sigma^2 \sum_{j=1}^{r} \lambda_j^{-1} t_j t_j^T \qquad (29)$$

The comparison of (28) and (29) reveals that the elimination of principal components will result in a decrease in the variance of parameters θ_{PC}. The diagonal elements of the matrix $T_s \Lambda_s^{-1} T_s^T$ are weighted sums of the inverse of the eigenvalues associated with the eliminated principal components. If these eigenvalues are small a substantial reduction in the variance of the PC estimates can be expected.

In practical application of the PC regression the problem of how many or if any principal components should be eliminated may arise. The answer to this problem could come from the diagnostic measures discussed and from commonly used least squares criteria as s^2, R^2 and others. Reference 5 presents the way for finding an optimal value of r based on the minimization of the criterion

$$(\tilde{\gamma}_{PC} - \gamma)^T (\tilde{\gamma}_{PC} - \gamma)$$

In the same reference it is also pointed out that the assumption of an integral rank for X can be too restrictive. A possible improvement to the principal components estimator, known as the fractional rank estimator, is introduced. If the rank of X lies in the interval (r, r + 1), the fractional rank estimator is given as

$$\tilde{\gamma}_{FR} = \left[\hat{\gamma}_1, \hat{\gamma}_2, \ldots, \hat{\gamma}_r, c\hat{\gamma}_{r+1}, 0 \ldots 0\right]^T$$

where the criterion for choosing c is given.

The second problem with the PC regression can be related to the formulation of regressors in equation (2). In numerical computation the measured data can enter the analysis in various forms. Probably the simplest approach would be to use regressors in their original form. If the scaled regressors to their unit length are preferred, different principal components will be obtained. The use of standardized regressors will result in X^TX being a matrix of correlations and the principal components will be changed again. This dependence of PC estimates on the form of regressors is obviously a weakness of this estimation technique.

MIXED ESTIMATION

The mixed estimation was developed as a Bayes-like technique by augmenting the measured data by prior information. For the linear model

$$Y = X\theta + \varepsilon \tag{2}$$

with $E(\varepsilon)$ and $E(\varepsilon^2) = \sigma^2 I$

it is assumed that $p < n$ prior restrictions on the elements of θ are available. These restrictions are formulated as

$$a = A\theta + \zeta \tag{30}$$

In (30) A is a matrix of each $p < n$ which includes known constants, a is a p-vector of values which can be specified, and ζ is a random vector with

$$E(\zeta) = 0,\ E(\zeta\varepsilon) = 0 \text{ and } E(\zeta^2) = \sigma^2 W$$

value W is a known weighting matrix.

Combining (2) and (30) the mixed model is given as

$$\begin{bmatrix} Y \\ a \end{bmatrix} = \begin{bmatrix} X \\ A \end{bmatrix} \theta + \begin{bmatrix} \varepsilon \\ \zeta \end{bmatrix} \tag{31}$$

For known σ^2 the application of least squares to (31) results in mixed estimation

$$\hat{\theta}_{ME} = (X^T X + A^T W^{-1} A)^{-1} (X^T Y + A^T W^{-1} a) \tag{32}$$

Introducing the augmented variables Y_a, X_a, and ε_a the mixed model can be also written as

$$Y_a = X_a \theta + \varepsilon_a \tag{33}$$

where, $E(\varepsilon_a) = 0$ and $E(\varepsilon_a^2) = \sigma^2 \begin{bmatrix} I & 0 \\ 0 & W \end{bmatrix} = \sigma^2 W_a$

Then, the mixed estimator can be expressed in the form

$$\hat{\theta}_{ME} = (X_a^T W_a^{-1} X_a)^{-1} X_a^T W_a^{-1} Y_a \tag{34}$$

It follows from the Gauss-Markoff theorem that $\hat{\theta}_{ME}$ given by (34) or (32) is an optimal unbiased linear estimator of θ.

In real application of the mixed estimation the a priori information is usually not known exactly. In this case

$$a = A\theta + b + \zeta \tag{35}$$

where $b \neq 0$ is an unknown vector. The mixed estimator corresponding to the condition given by (35) will be called $\tilde{\theta}_{ME}$. The expected value of $\tilde{\theta}_{ME}$ is obtained by substituting (2) and (35) into (32)

$$\begin{aligned} E(\tilde{\theta}_{ME}) &= E[M^{-1}X^T X\theta + M^{-1}X^T\varepsilon + M^{-1}A^T W^{-1}A\theta \\ &\quad + M^{-1}A^T W^{-1}b + M^{-1}A^T W^{-1}\zeta] \qquad (36) \\ &= \theta + M^{-1}A^T W^{-1}b \end{aligned}$$

where $$M = X^T X + A^T W^{-1} A. \tag{37}$$

The estimate $\tilde{\theta}_{ME}$ is therefore biased by the quantity $M^{-1}A^T W^{-1}b$.

The covariance matrix of the mixed estimator is

$$\text{Cov}(\tilde{\theta}_{ME}) = \sigma^2 M^{-1} \tag{38}$$

The difference between the covariance of the LS and ME is given as

$$\text{Cov}(\hat{\theta}) - \text{Cov}(\tilde{\theta}_{ME}) = \sigma^2(X^T X)^{-1} - \sigma^2(X^T X + A^T W^{-1} A)^{-1} \tag{39}$$

Since $\text{Cov}^{-1}(\tilde{\theta}_{ME}) - \text{Cov}^{-1}(\hat{\theta}) = \sigma^{-2}(A^T W^{-1} A) \geq 0$, the right side of (39) is a nonnegative definite matrix. This means that the addition of the priori information to the ordinary regression will result in reduction of variance of the LS estimates.

The restrictions on parameters given by (30) can take several forms. The most common are:

a) a separate estimate of all parameters θ_j, $j = 1, 2, \ldots, n$ exists. If these estimates are called θ_0, then (30) is changed as

$$\theta_0 = \theta + \zeta$$

which means that $a = \theta_0$ and A is the m x n identity matrix. If only the estimates of some elements of the vector θ are known a priori, say $\theta_{0,1}$, then

$$\theta_{0,1} = \theta_j + \zeta$$

and $A = [I, 0]$ where the dimension of I in A corresponds to the dimension of θ_j.

b) Special case occurs when $W = 0$. This corresponds to knowing that $a = A\theta$ with certainty. This situation leads to a piecewise regression discussed in ref. 6.

c) Sometimes the a priori information is given as a statement that particular parameters lie in a certain region (d_{min}, d_{max}). For the parameter θ_j it means that

$$\theta_{0,j} = 1/2\ (d_{min} + d_{max}) + \zeta_j$$

EXAMPLES

The detection of collinearity and the biased estimation techniques described are demonstrated in two examples. The flight test data for these examples were obtained from the longitudinal small-amplitude maneuvers of a highly augmented, inherently unstable research aircraft. The longitudinal motion of this aircraft was controlled by three surfaces, canard, flaperons

and strake, moved by an automatic control system. The data used in the analysis were in the form of sampled time histories of open-loop input variables δ_c, δ_f and δ_s and output variables of V, α, q, a_z and $\dot{q}$. The model for the vertical-force and pitching-moment coefficient was formulated as

$$C_a = C_{a_0} + C_{a_\alpha}\alpha + C_{a_q}\frac{q\bar{c}}{2V} + C_{a_{\delta s}}\delta_s + C_{a_{\delta f}}\delta_f + C_{a_{\delta c}}\delta_c \qquad (40)$$

for a = Z or m. In (40) the regressors are represented by the increments of the input and output variables from their values in steady flight conditions prior to the excited motion. The independent variables in (40) were computed from the expressions

$$C_Z = \frac{mg}{\bar{q}S}\, a_z$$

$$C_m = \frac{I_Y}{\bar{q}S\bar{c}}\,\dot{q}$$

The unknown parameters in (39) are the stability and control derivatives, and the bias term C_{a_0}.

Example 1. Three control variables:

The aircraft short-period response to a series of commanded pitch doublets is illustrated in figure 2. Shown are time histories of three longitudinal control and three output variables. Inspection of figure 2 reveals very close relationship among all three open-loop inputs, thus indicating strong possibility for data collinearity. For the assessment of collinearity the correlation matrix of standardized regressors was formulated and its determinant computed. The correlation matrix is shown in Table I. By examining this matrix the simple correlation greater than .80 between two pairs of regressors (δ_c, α) and (δ_c, δ_s) can be seen. The determinant value was found equal to 0.00106. Therefore, the high pairwise correlations and the low value of the determinant point out data collinearity. Because of the weakness of the VIF as a diagnostic measure its values are not given.

In order to decide which regressors are affected by collinearity the variance proportions were computed. They are presented in Table II for scaled regressors in (40). Also included in the table are the eigenvalues of the X^TX matrix and condition numbers. The variance proportions corresponding to the largest condition number indicate four damaging dependencies involving δ_c, δ_s, q and the bias term. The second dependency involves α and δ_f. It corresponds to the condition number κ = 36 which may be considered too small for having any serious effect on the estimates.

As the result of data collinearity assessment it was decided to use the principal components regression with the smallest eigenvalue of X^TX equal to zero, thus reducing rank of the X matrix by one (r = 5). For the mixed estimation the parameters $C_{Z_{\delta s}}$ and $C_{m_{\delta s}}$ were set at their wind-tunnel values with the uncertainties estimated from repeated measurements in different facilities and for different configurations. The selection of strake terms was based on small sensitivity of these parameters and expected sufficient accuracy of their a priori values. The a priori values and three different values of their variance used in the mixed estimation are given in Table III.

In Table IV the results of the least squares, principal components and mixed estimation of parameters in the equation for C_Z are summarized. Presented are the mean values and standard errors of parameter estimates, and the standard errors of the C_Z estimates using the residuals. Also included are the sensitivities computed for wind-tunnel values given in the last column of the table, and the increments of the squared multiple correlation coefficient due to regressors in (40). Both the sensitivity analysis and squared multiple correlation coefficients indicate that the only important term in the equation for C_Z is $C_{Z_\alpha}\alpha$. This term in combination with C_{Z_0}, explain 99% of variation in the measured data. It can be, therefore, expected that data collinearity combined with low sensitivities will cause severe identifiability problems for most of the other parameters. These problems are immediately apparent from the LS estimates of $C_{Z_{\delta s}}$ and C_{Z_q} which are much higher than that from the wind tunnel and theory respectively.

The principal components regression was first applied to scaled data. The results show no improvement over the LS results. When the original regressors were used, however, the parameters $C_{Z_{\delta s}}$ and $C_{Z_{\delta c}}$ came out with correct sign. There was a small increase in C_{Z_α} but substantial decrease in C_{Z_q}. The fit to the data, measured by the standard error of C_Z, deteriorated. No explanation has been found for the differences between the two sets of principal components estimates. The mixed estimation with moderate and tight restrictions on the a priori value gave the best sets of estimates when compared with the wind-tunnel data and results of the two previous techniques.

The results from the data governed by the pitching-moment equation are presented in Table V. In the model for C_m two terms, $C_{m_{\delta c}}\delta_c$ and $C_{m_\alpha}\alpha$, are important. Together they explain 97% of the variation in the data. The principal components regression with the original regressors improves the LS estimates of $C_{m_{\delta c}}$ and $C_{m_{\delta s}}$ and makes them consistent with the mixed estimates under moderate or tight restrictions. A serious problem with the principal components regression is the non-physical value for the parameter C_{m_q}.

Example 2. Two control variables:

The second maneuver was commanded by two pitch doublets. This time the control system held the flaperons at constant deflection. The time histories of input variables δ_c and δ_s, and output variables α, q, and a_z are plotted in figure 3. The correlation matrix is given in Table VI showing strong correlation between (α, δ_s) and (δ_s, δ_c). The value of the determinant was equal to 0.0100. The various proportions, eigenvalues of X^TX and condition numbers are presented in Table VIII. The damaging dependencies for the largest condition number are again among the regressors δ_s, δ_c, q and the bias term.

Table VIII and IX present the estimation results, sensitivities, squared multiple correlation coefficients and wind-tunnel data used in the mixed estimation (values for $C_{Z_{\delta s}}$ and $C_{m_{\delta s}}$) and for comparison. The principal components regression using the original data with $r = 4$ did not bring the expected improvement over the LS estimates. Therefore further reduction in rank of the X matrix was attempted. The following estimates of the important parameters C_{Z_α}, C_{m_α} and $C_{m_{\delta c}}$ were much closer to those from the mixed estimation. By observing the results for $r = 3$ and $r = 4$, and the results from the mixed estimation, it is possible to argue that the optimum r should be somewhere between 3 and 4. Such selection of r would, however, lead to the fraction rank estimator mentioned earlier but not developed in this report.

As in the previous example, the mixed estimation with moderate or tight restrictions (see Table III) resulted in the best sets of estimates. The technique only failed to provide some physical values for the damping terms C_{Z_q} and C_{m_q}. This failure could be explained by very low sensitivities of these parameters.

CONCLUDING REMARKS

Near linear dependency in measured data, called collinearity, and its effect on linear regression were briefly discussed. Then, procedures for detection and assessment of collinearity were presented. They included the evaluation of the correlation matrix and its inverse, eigensystem analysis or singular value decomposition, and parameter variance decomposition. The first of these procedures is relatively simple and straight forward but, it cannot reveal the presence of several coexisting dependencies among the regressors. Eigensystem analysis examines the values of eigenvalues in the matrix composed

by regressors. The large condition numbers serve as indicators of data collinearity. The singular value decomposition provides similar information. It is preferred by some analysts because of greater stability in its computation. Both approaches become more effective when combined with parameter variance decomposition. This combination can find which regressors are near linearly dependent and indicates what action should be taken in order to lessen the effect of collinearity on the estimates. In connection with data collinearity the problem of parameter identification was also addressed and the sensitivity analysis as a tool for its assessment introduced.

One way of dealing with collinearity is to use different estimation techniques from the ordinary least squares. This report explained the reasons for using the biased estimation techniques and presented two of these techniques, principal components regression and mixed estimation. The principal components regression eliminates the effect of small eigenvalues by reducing rank of the matrix of regressors. The weakness of this technique can be seen in the restriction to an integral rank and the dependence of the estimates on various forms of the regressors (original, scaled or standardized). The mixed estimation is a Bayes-like technique which is applied to measured data augmented by prior information. This estimation procedure can be very successful provided that a priori values of selected parameters are known with reasonable accuracy.

The detection and assessment of collinearity, and the two biased estimation techniques were demonstrated in two examples using flight data from longitudinal maneuvers of an experimental aircraft. In these examples the correlation matrix of regressors indicated the existence of correlation between two pairs of regressors. The variance proportions, however, determined which regressors were affected by collinearity. The estimates of parameters in the aerodynamic model equations for the vertical-force and pitching-moment coefficient were also obtained by the ordinary least squares. These results confirmed a damaging effect of collinearity on the estimated values and their standard errors. The principal components regression provided substantially improved estimates with the exception of damping-in-pitch derivative. Some further improvement was obtained from mixed estimation where the a priori values were taken from wind-tunnel data. The parameter estimates were completed by the results of the sensitivity analysis and by increments in the squared multiple correlation coefficient indicating the importance of individual terms in regression equations. The proposed procedure for dealing with data collinearity proved that it could become a useful approach for estimating parameters of a highly augmented, possibly unstable aircraft from flight data.

REFERENCES

1. Montgomery, Douglas C. and Peck, Elizabeth A.: "Introduction to Linear Regression Analysis," John Wiley & Sons, Inc., 1982.

2. Belsley, David A.; Kuh, Edwin; and Welsch, Roy E.: "Regression Diagnostics: Identifying Influential Data and Sources of Collinearity," John Wiley & Sons, Inc., 1980.

3. Theil, Henri: "Principles of Econometric," John Wiley & Sons, Inc., 1971.

4. Hsia, T. C.: "System Identification, Least-Squares Methods," Lexington Books, 1977.

5. Hocking, R. R.; Speed, F. M.; and Lynn, M. J.: "A Class of Biased Estimators in Linear Regression," Technometrics, Vol. 18, No. 4, November 1976, pp. 425-437.

6. Toutenburg, Helge: "Prior Information in Linear Models," John Wiley & Sons, Inc., 1982.

APPENDIX A

SCALED AND STANDARDIZED REGRESSORS

In order to have the columns of the X matrix of unit length, the original regressors x_j are replaced by scaled regressors x'_j using the formula

$$x'_{ji} = \frac{x_{ji}}{\sum_{i=1}^{N} x_{ji}^2} \tag{A.1}$$

for $j = 1, 2, \ldots, n$, and $i = 1, 2, \ldots, N$.

Using the scaled regressor the model in (2) is changed as

$$Y = X' \theta' \tag{A.2}$$

where

$$X' = \begin{bmatrix} \frac{1}{\sqrt{N}} & x'_{11} & x'_{21} & \cdots & x'_{n1} \\ \frac{1}{\sqrt{N}} & x'_{12} & x'_{22} & \cdots & x'_{n2} \\ \frac{1}{\sqrt{N}} & x'_{1N} & x'_{2N} & \cdots & x'_{nN} \end{bmatrix} \tag{A.3}$$

and the new parameters θ'_j are related to original parameters θ_j by the equation

$$\theta_j = \frac{\theta'_j}{\sum_{i=1}^{N} x_{ji}^2} \tag{A.4}$$

The standardized (scaled and centered) regressors x^*_j are obtained as

$$x^*_{ji} = \frac{x_{ji} - \bar{x}_j}{S_j} \tag{A.5}$$

where $\bar{x}_j = \frac{1}{N} \sum_{i=1}^{N} x_{ji}$

$$S_j = \sum_{i=1}^{N} (x_{ji} - \bar{x}_j)^2$$

The regression model has the form

$$Y = X^* \theta^* \tag{A.6}$$

with the LS estimates

$$\hat{\theta}^* = (X^{*T} X^*)^{-1} X^{*T} Y \tag{A.7}$$

where $(X^{*T} X^*)$ is the correlation matrix of regressors. The original parameters are related to the parameters in (A.6) as

$$\theta_j = \frac{\theta^*_j}{S_j} \tag{A.8}$$

and $$\theta_0 = \theta^*_0 - \frac{\theta^*_1}{S_1} \bar{x}_1 - \frac{\theta^*_2}{S_2} \bar{x}_2 - \ldots - \frac{\theta^*_n}{S_n} \bar{x}_n \tag{A.9}$$

APPENDIX B

SQUARED MULTIPLE CORRELATION COEFFICIENT

The regression equation with LS estimates is given as

$$Y = X \hat{\theta} + \hat{\varepsilon} \tag{B.1}$$

where it is assumed that the regressors and dependent variable are centered. Premultiplying each side of (B.1) by its own transpose results in

$$Y^T Y = \hat{\theta}^T X^T X \hat{\theta} + \hat{\varepsilon}^T \hat{\varepsilon} \tag{B.2}$$

The term $X^T \hat{\varepsilon} = 0$ because the vector of residuals $\hat{\varepsilon}$ is orthogonal to each of the n columns of X. From (B.2) it can be concluded that a fraction R^2 of $\sum_{i=1}^{N} y_i^2$ is accounted for the regressors and that a fraction $1 - R^2$ is represented by residuals. Then

$$R^2 = \frac{\hat{\theta}^T X^T X \hat{\theta}}{Y^T Y}$$

and (B.3)

$$1 - R^2 = \frac{\hat{e}^T \hat{e}}{Y^T Y}$$

The R^2 is known as the squared multiple correlation coefficient associated with (B.1). This coefficient can be interpreted as a measure of variability in y explained by the regression model.

For the regression equation with the bias term θ_0 the measure $\sum_{i=1}^{N} y_i^2$ is replaced by the sum of squared values taken as deviations from the mean, i.e. $\sum_{i=1}^{N} (y_i - \bar{y})^2$. With the new measure the expressions for R^2 will take the form

$$R^2 = \frac{\hat{\theta}^T X^T X \hat{\theta} - N\bar{y}^2}{Y^T Y - N\bar{y}^2} \qquad (B.4)$$

$$1 - R^2 = \frac{\hat{\varepsilon}^T \hat{\varepsilon}}{Y^T Y - N\bar{y}^2}$$

where $\bar{y} = \frac{1}{N} \sum_{i=1}^{N} y_i$.

TABLE I. - $X^{*T}X^{*}$ MATRIX IN CORRELATION FORM
EXAMPLE 1.

α	$q\bar{c}/2V$	δ_s	δ_f	δ_c
1.000	.316	.745	.682	- .839
	1.000	.499	- .287	- .134
		1.000	.481	- .891
			1.000	- .794
				1.000

TABLE II. - COLLINEARITY DIAGNOSTIC FOR
DATA IN EXAMPLE 1.

Eigen-values	Condition number	Variance proportions (scaled regressors)					
		1	α	$q\bar{c}/2V$	δ_s	δ_f	δ_c
3.2401	1	.000	.018	.032	.001	.003	.000
1.4295	2	.005	.000	.000	.031	.018	.000
.9823	3	.001	.111	.047	.002	.027	.001
.2487	13	.026	.103	.002	.106	.008	.020
.0908	36	.053	.755	.010	.014	.739	.005
.0086	377	.914	.014	.909	.845	.205	.974

TABLE III. - A PRIORI DATA. EXAMPLE 1.

A priori	$(C_{Z_{\delta e}})_0 = -.20 + \zeta$		$(C_{m_{\delta e}})_0 = -.33 + \zeta$	
	95% range	Implied variance	95% range	Implied variance
Loose	- .28, - .12	.0016	- .45, - .21	.0036
Medium	- .24, - .16	.0004	- .41, - .25	.0016
Tight	- .22, - .18	.0001	- .37, - .29	.0004

TABLE IV. - LEAST SQUARES AND BIASED ESTIMATES, AND WIND-TUNNEL DATA FOR VERTICAL-FORCE PARAMETERS. EXAMPLE 1.

Parameter	Sensitivity	Least squares		Principal components (r = 5)		Mixed estimation			Wind Tunnel
						Variance			
		$\hat{0}$	ΔR^2 [%]	Scaled regressors	Original regressors	.0016	.0004	.0001	
C_{Z_0}		- .2710 (.00017)		- .2710 (.00016)	- .2715 (.00022)	- .2708 (.00017)	- .2707 (.00017)	- .2707 (.00017)	
C_{Z_α}	3.0088	- 4.67 (.022)	99.01	- 4.69 (.018)	- 4.99 (.021)	- 4.75 (.021)	- 4.79 (.020)	- 4.81 (.019)	- 5.22
C_{Z_q}	.0012	- 34 (1.7)	.11	- 32 (1.1)	- .063 (.0012)	- 26 (1.4)	- 21 (1.2)	- 19 (1.1)	- 8[a]
$C_{Z_{\delta s}}$	.0076	.29 (.042)	.05	.24 (.0080)	- .38 (.033)	.03 (.029)	- .11 (.018)	- .17 (.0097)	- .20
$C_{Z_{\delta f}}$	.0859	- 1.46 (.038)	.70	- 1.49 (.030)	- 1.52 (.050)	- 1.59 (.035)	- 1.66 (.033)	- 1.69 (.032)	- 1.28
$C_{Z_{\delta c}}$	.0278	.11 (.037)	.01	.66 (.0036)	- .43 (.034)	- .11 (.027)	- .23 (.019)	- .29 (.013)	- .27
$s(C_Z)$		.00283		.00283	.00374	.00293	.00306	.00314	

[a]Theoretical value; Note: the values in parenthesis are standard errors

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE V. - LEAST SQUARES AND BIASED ESTIMATES, AND WIND-TUNNEL DATA FOR PITCHING-MOMENT PARAMETERS. EXAMPLE 1.

Parameter	Sensitivity	Least squares		Principal components (r = 5)		Mixed estimation			Wind Tunnel
						Variance			
		$\hat{\theta}$	ΔR^2 [%]	Scaled regressors	Original regressors	.0036	.0016	.0004	
C_{m_0}		.0003 (.00022)		.0009 (.00021)	.0001 (.00022)	.0003 (.00022)	.0004 (.00022)	.0004 (.00022)	
C_{m_α}	.3080	1.54 (.029)	15.19	1.32 (.023)	1.46 (.021)	1.52 (.027)	1.51 (.026)	1.50 (.025)	1.67
C_{m_q}	.0012	- 10 2.2)	.48	11 (1.4)	.016 (.0013)	- 8 (1.9)	- 6 (1.7)	- 5 (1.5)	- 8[a]
$C_{m_{\delta s}}$	.0206	- .14 (.054)	.03	- .80 (.010)	- .34 (.033)	- .23 (.040)	- .26 (.032)	- .31 (.019)	- .33
$C_{m_{\delta f}}$	.0159	- .57 (.050)	.24	- .96 (.039)	- .59 (.050)	- .62 (.046)	- .63 (.045)	- .66 (.043)	- .55
$C_{m_{\delta c}}$	.2949	1.03 (.049)	82.30	.432 (.0047)	.87 (.034)	.96 (.037)	.92 (.031)	.89 (.021)	.88
$s(C_m)$		.00367		.00417	.00374	.00368	.00368	.00370	

[a]Theoretical value; Note: the values in parenthesis are standard errors

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE VI. - $X^{*T}X^{*}$ MATRIX IN CORRELATION FORM. EXAMPLE 2.

α	$q\bar{c}/2V$	δ_e	δ_c
1.000	.289	.813	- .780
	1.000	.227	- .159
		1.000	- .987
			1.000

TABLE VII. - COLLINEARITY DIAGNOSTIC FOR DATA IN EXAMPLE 2.

Eigen-values	Condition number	Variance proportions (scaled regressors)				
		1	α	$q\bar{c}/2V$	δ_e	δ_c
2.7482	1	.000	.494	.000	.003	.000
1.0916	3	.011	.095	.000	.049	.000
.9184	3	.002	.003	.204	.001	.000
.2268	12	.063	.204	.012	.031	.073
.0150	184	.924	.203	.784	.917	.926

TABLE VIII. - LEAST SQUARES AND BIASED ESTIMATES, AND WIND-TUNNEL DATA FOR VERTICAL-FORCE PARAMETERS. EXAMPLE 2.

Parameter	Sensitivity	Least squares		Principal components		Mixed estimation			Wind Tunnel
						Variance			
		$\hat{\beta}$	ΔR^2 [%]	r = 4	r = 3	.0016	.0004	.0001	
C_{Z_0}		- .1514 (.00033)		- .1505 (.00038)	- .1532 (.00033)	- .1527 (.00025)	- .1528 (.00025)	- .1528 (.00025)	
C_{Z_α}	.77777	- .47 (.042)	97.97	- 5.49 (.050)	- 5.68 (.060)	- 5.54 (.040)	- 5.54 (.040)	- 5.54 (.040)	- 5.90
C_{Z_q}	.00006	- 36 (3.4)	.61	- .09 (.0059)	- .02 (.00021)	- 47 (3.0)	- 47 (2.9)	- 48 (2.9)	- 7[a]
$C_{Z_{\delta s}}$	.00023	- 1.2 (.16)	.11	- 2.0 (.16)	- .11 (.020)	- .20 (.039)	- .16 (.020)	- .14 (.010)	- .14
$C_{Z_{\delta c}}$	.00546	- 1.3 (.11)	.81	- 2.0 (.11)	- .66 (.036)	- .70 (.042)	- .67 (.035)	- .66 (.033)	- .56
$s(C_Z)$		.00352		.00426	.00538	.00377	.00380	.00381	

[a]Theoretical value; Note: the values in parenthesis are standard errors

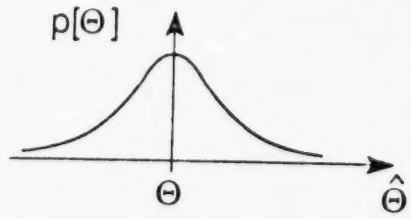
ORIGINAL PAGE IS
OF POOR QUALITY

TABLE IX. - LEAST SQUARES AND BIASED ESTIMATES, AND WIND-TUNNEL DATA FOR PITCHING-MOMENT PARAMETERS. EXAMPLE 2.

Parameter	Sensitivity	Least squares		Principal components		Mixed estimation			Wind Tunnel
						Variance			
		$\hat{\beta}$	ΔR^2 [%]	r = 4	r = 3	.0036	.0016	.0004	
C_{m_0}		.0003 (.00013)		.0002 (.00017)	.0002 (.00018)	.0010 (.00015)	.0011 (.00014)	.0013 (.00014)	
C_{m_α}	.00804	.32 (.023)	6.96	.32 (.0023)	.42 (.0028)	.36 (.022)	.37 (.022)	.38 (.022)	.60
C_{m_q}	.00012	2 (1.9)	.11	.036 (.0027)	.0017 (.00010)	9 (1.6)	10 1.6)	11 (1.6)	- 10[a]
$C_{m_{\delta s}}$	.00098	.49 (.086)	1.47	.55 (.074)	- .395 (.0094)	- .35 (.049)	- .15 (.036)	- .25 (.019)	- .29
$C_{m_{\delta c}}$	.01742	1.24 (.061)	89.94	1.29 (.052)	.64 (.017)	.89 (.038)	.81 (.030)	.74 (.022)	.78
$s(C_m)$		.00191		.00192	.00251	.00206	.00213	.00219	

[a]Theoretical value; Note: the values in parenthesis are standard errors

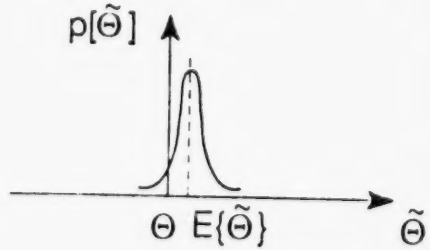
Unbiased estimator



$E\{\hat{\Theta}\} = \Theta$ ... unbiased

Var $\{\hat{\Theta}\}$ ... large if information matrix ill-conditioned

Biased estimator



$E\{\hat{\Theta}\} = \Theta$ ... biased

Var $\{\hat{\Theta}\}$ ... small

Figure 1. - Distributions of unbiased and biased estimators of .

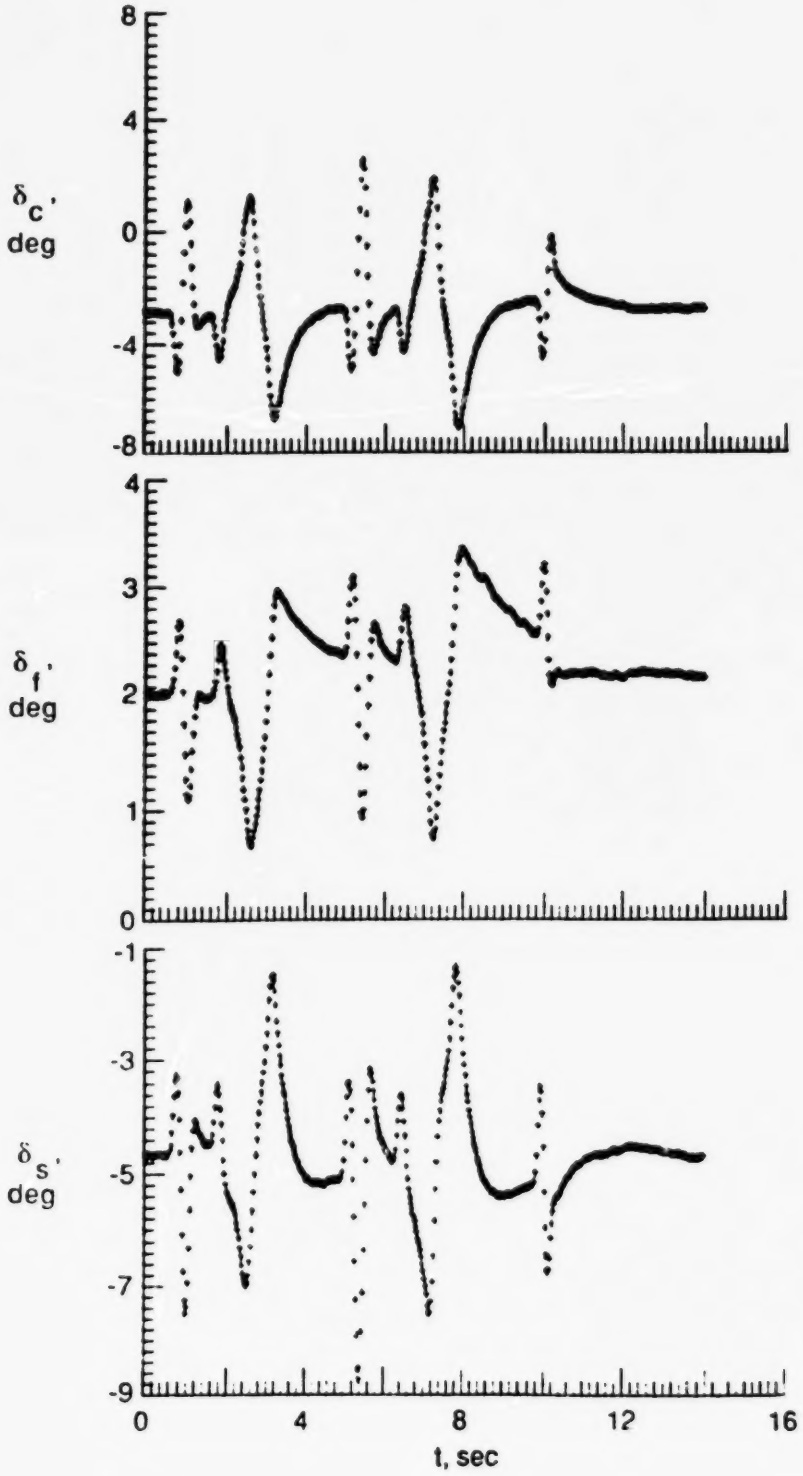


Figure 2. - Time histories of measured longitudinal variables. Example 1.

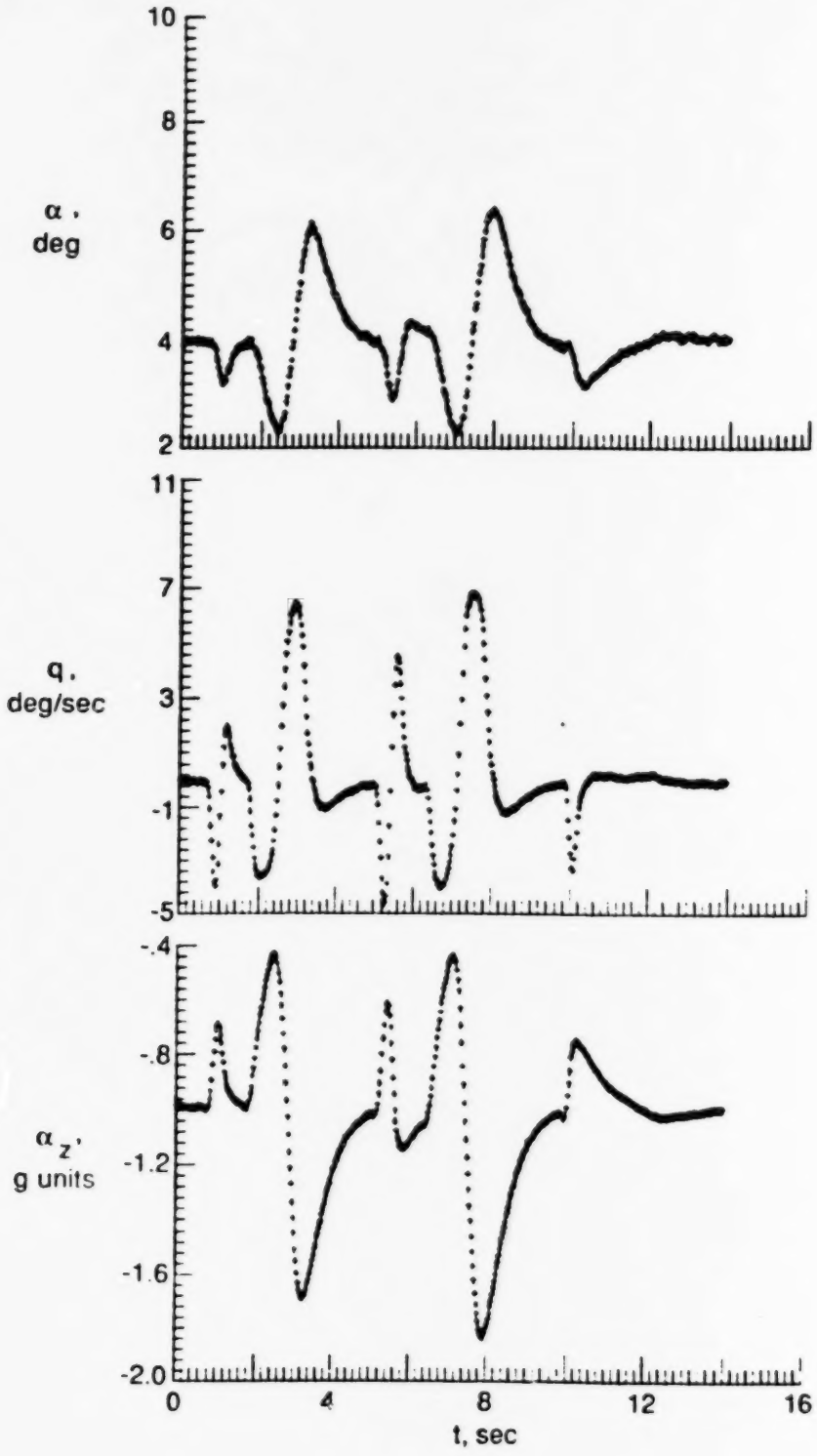


Figure 2. - Concluded

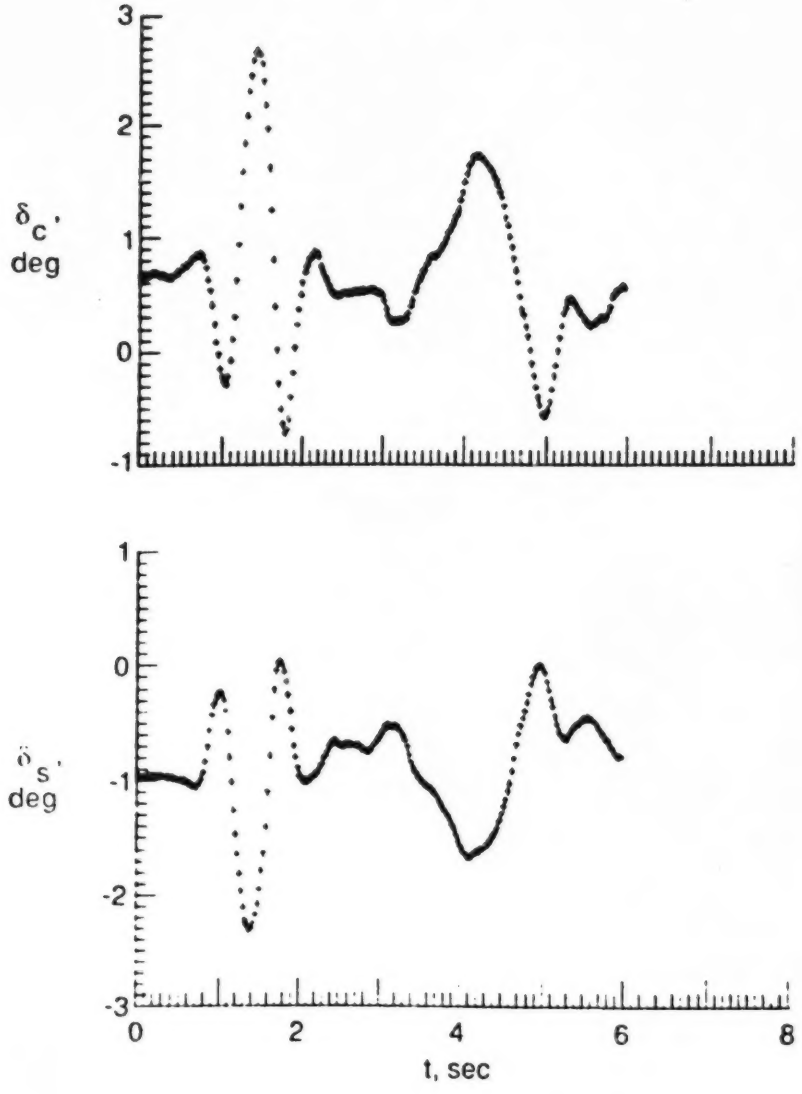


Figure 3. - Time histories of measured longitudinal variables. Example 2.

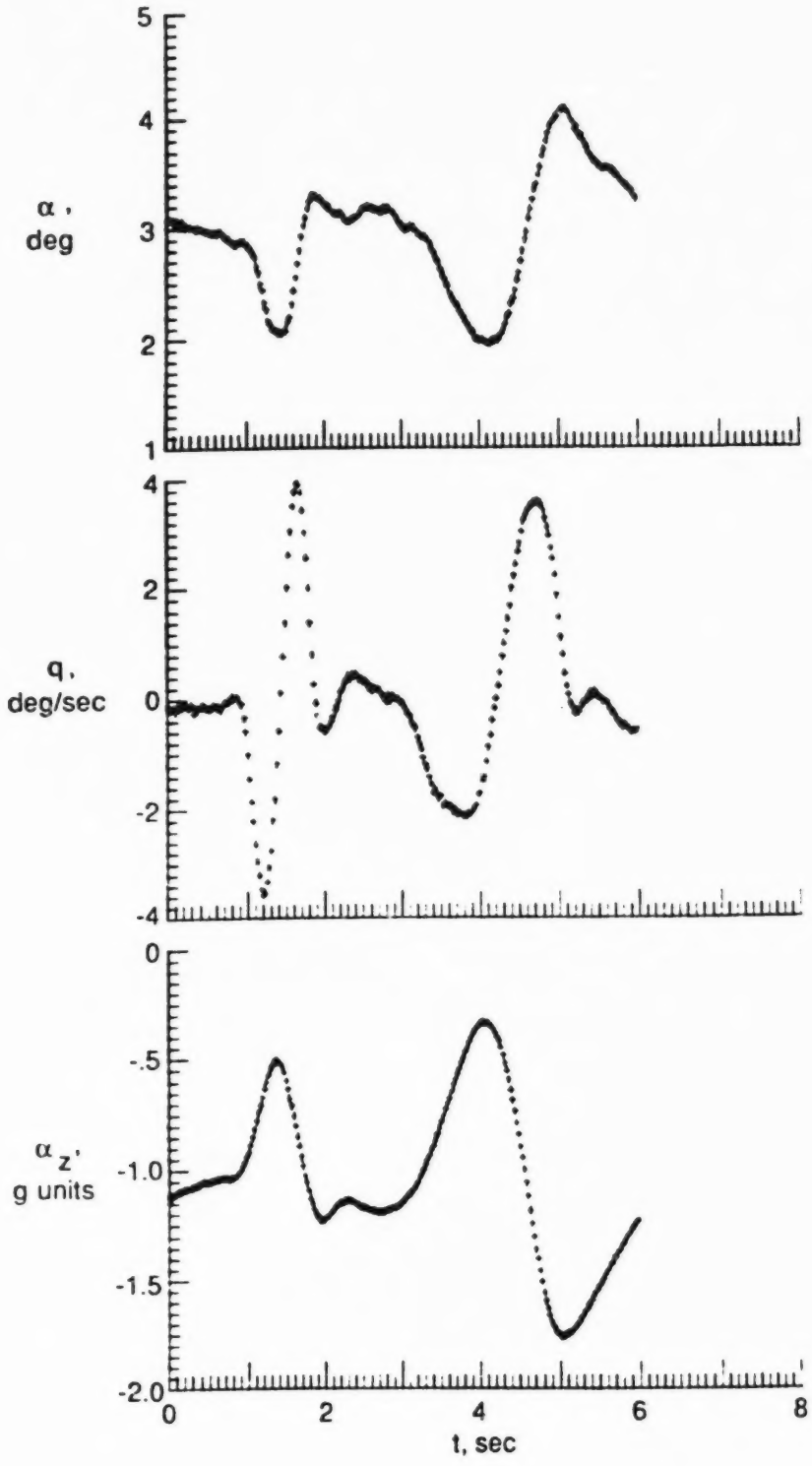


Figure 3. - Concluded

ORIGINAL PAGE IS
OF POOR QUALITY

NASA

Report Documentation Page

1 Report No	2 Government Accession No	3 Recipient's Catalog No
NASA TM-100649		

4 Title and Subtitle

Two Biased Estimation Techniques in Linear Regression--Application to Aircraft

5 Report Date

July 1988

6 Performing Organization Code

7 Author(s)

Vladislav Klein

8 Performing Organization Report No

10 Work Unit No.

505-66-01-02

9 Performing Organization Name and Address

NASA Langley Research Center
Hampton, VA 23665-5225

11 Contract or Grant No

13 Type of Report and Period Covered

Technical Memorandum

12 Sponsoring Agency Name and Address

National Aeronautics and Space Administration
Washington, DC 20546-0001

14 Sponsoring Agency Code

15 Supplementary Notes

Vladislav Klein: The George Washington University, Joint Institute for Advancement of Flight Sciences, Hampton, Virginia.

16 Abstract

Several ways for detection and assessment of collinearity in measured data are discussed. Because data collinearity usually results in poor least squares estimates, two estimation techniques which can limit a damaging effect of collinearity are presented. These two techniques, the principal components regression and mixed estimation, belong to a class of biased estimation techniques. Detection and assessment of data collinearity and the two biased estimation techniques are demonstrated in two examples using flight test data from longitudinal maneuvers of an experimental aircraft. The eigensystem analysis and parameter variance decomposition appeared to be a promising tool for collinearity evaluation. The biased estimators had far better accuracy than the results from the ordinary least squares technique.

17 Key Words (Suggested by Author(s))	18 Distribution Statement
Data collinearity Linear regression Biased estimation Flight test data analysis	Unclassified - Unlimited Subject Category 65

19 Security Classif (of this report)	20 Security Classif (of this page)	21 No of pages	22 Price
Unclassified	Unclassified	30	A03

For sale by the National Technical Information Service, Springfield, Virginia 22161-2171